Defense Technical Information Center Compilation Part Notice

ADP014610

TITLE: Comparison of Approaches to Obtaining a Transformation Matrix Effecting a Fit to a Factor Solution Obtained in a Different Sample

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Proceedings of the Eighth Conference on the Design of Experiments in Army Research Development and Testing

To order the complete compilation report, use: ADA419759

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP014598 thru ADP014630

COMPARISON OF APPROACHES TO OBTAINING A TRANSFORMATION MATRIX EFFECTING A FIT TO A FACTOR SOLUTION OBTAINED IN A DIFFERENT SAMPLE

Cecil D. Johnson
U. S. Army Personnel Research Office

BACKGROUND. The importance of physical proficiency measures to the selection and evaluation of Army personnel can scarcely be questioned. Examination of the duty assignments prevalent in various Army jobs indicates clearly that physical strength, endurance and coordination are often highly important factors in job success. At the United States Military Academy in particular, considerable attention has been devoted to physical training and to the measurement of various aspects of physical abilities or physical proficiency among cadets at West Point. Various tests of physical proficiency were introduced in the physical aptitude entrance examination procedure or studied for possible use.. They have been examined both as individual measures and as component parts of various batteries. Several factor analyses of large batteries of physical proficiency measures, physical education, grades, and other variables were accomplished in previous studies. These studies had as their objective the identification of basic underlying physical ability variables that possess the simplifying statistical characteristics frequently referred to as simple structure. These basic variables, or factors, aid in understanding the nature of the scores, in eliminating duplicating measures, and in suggesting new tests.

The several factor solutions available for comparison contain numerous variables in common, other similar variables (as when a 25-yard dash is substituted for a 30-yard dash), and still other variables which are unique for a particular solution. This paper considers several methods for comparing solutions obtained in these separate studies involving physical proficiency and related measures.

The problem of approximating in a second sample, a rotated factor solution originally obtained in a previously analyzed sample is also present in another Army Personnel Research Office research study currently in the final computing phase. This study involved thirty-one psychological tests. Some of these tests are measures of intellectual ability, others are measures of cognitive information, and others are non-cognitive measures in the "personality" domain. As is typical with Personnel Research Office factor analysis studies, the objective was the identification of constructs which would predict the performance of soldiers on the job. In this case the job was that of an enlisted Infantryman and the measure of performance was obtained from ratings by superiors and peers at the close of maneuvers in

Germany. The tests had been administered to the enlisted men on their entry into the Army.

An initial principal component factor solution was transformed by an orthogonal matrix so as to provide simple structure. The initial factor solution can be described as a matrix whose elements are the correlations between the tests and standard length orthogonal reference vectors. This solution usually provides parsimony in that a relatively small number of reference vectors is needed to closely approximate the test correlation matrix when the factor matrix is post-multiplied by its transpose. However, the psychologist wants the reference vectors, or factors, to have additional properties implied by the concept of a simple structure. If simple structure is present among the reference vectors, each reference vector has high correlations with a few tests and approximately zero correlations with the remainder. Furthermore, the tests with which a particular reference factor has a high relationship will be relatively independent of the other reference vectors. It is apparent that the presence of simple structure permits the psychologist to interpret the reference vectors in terms of his test, and if the orthogonality of the reference vectors is retained, as when the transformation matrix is orthogonal, all the original parsimony of the initial principal component factor solution is retained. Psychologists usually refer to the process of transforming a solution to simple structure as rotation, and call the transformed solution a "rotated" solution.

In the factor analysis of psychological tests described above, the rotated solution, when extended to the rating variables, displayed a very interesting relationship between the rotated factors and the performance measures. One cognitive factor and one non-cognitive factor predicted performance while all other factors had a zero relationship with performance. It became a matter of considerable interest to determine whether these relationships could be verified in an independent sample where properties of the sample had not been used to determine the particular transformation used to obtain the rotated factor solution. Both factors retained their validity in the cross (independent) sample, but an additional factor (previously non-valid) also displayed a smaller amount of validity. In this study the factor validities in the first sample were fairly well replicated in the second sample.

Thus, both studies, the one involving physical proficiency variables and the one involving psychological tests, require an initial factor solution in a second independent sample, the transformation of this solution to one

approximating the rotated solution in the first sample, and finally, the extension* of the transformed solution to non-overlapping predictor variables and/or criterion variables.. The criterion variables may well overlap across the two studies but should be withheld from the initial factor analysis for two reasons:

(1) It is desirable that the factors be entirely defined by predictor variables.

(2) The validity of the transformed factors are being determined in the independent, or cross sample. Thus, the definition of the factors in the cross sample must be independent of the criterion variables.

B. F. Green has reported a method for computing an orthogonal transformation matrix which will minimize the sum of squares of the differences between the transformed matrix and the matrix to be fitted. However, his derivation does not generalize so as to provide an orthonormal transformation that can utilize more reference vectors in the cross sample than are in the matrix to be fitted.

If the investigator is confident that the initial cross sample factor solution does not have a rank which exceeds the rank of the solution to be fitted, this orthogonal transformation is clearly suitable. On the other hand, if in the cross sample there is likely to be considerable variance common to two or three variables that is not explained by the more general common factors utilized in the initial sample, the advantages of an orthonormal solution become apparent.

APPROACH AND RESULTS. Thus, in obtaining the transformation matrix necessary for fitting K factors, the investigator has a choice of using a method which obtains the best orthogonal transformation matrix applicable to the first K factors, or he can choose to use a non-square orthonormal transformation matrix which can be applied to a full factorization, i.e., to as many factors as there are variables. The first method, using an orthogonal transformation, requires the fitted solution to reproduce

---

*This factor extension is accomplished by post multiplying the m x n matrix of correlation coefficients (between the overlapping and non-overlapping) by $A_y D_y^{-1/2}$, where R_y is the matrix of correlation coefficients among the n overlapping variables and $A'_y R_y A_y = D_y$; $A'_y = A_y^{-1}$; D_y = eigen values.

the cross sample correlation matrix to the full extent possible with a principal component solution. The second method, using an orthonormal transformation, provides a less exact reproduction of the correlation matrix, but permits a better fit to the reference factors -- if the dimensionality of the experimental variable space exceeds the number of factors.

Since in the physical proficiency study the number of common variables, for some of the comparisons, was large as compared to the number of factors being fitted, and the initial solutions had been obtained using a correlation matrix involving even more variables, the non-square, orthonormal transformation was utilized. A description of this technique is provided in the hand-out.

The three rotated factor solutions to which the initial solutions in the cross-samples were fitted, tended to have smaller communalities than the four cross-sample fitted solutions. This is possibly explained by two things. The methods of obtaining the initial solutions (that were subsequently rotated in the reference samples) were less efficient that the principle component method used for the initial cross sample solutions. Also, the initial factor solutions in the first sample were obtained to span the non-common variables as well, whereas the cross sample solutions were obtained on the common variables only.

In comparing the use of the orthonormal as compared to an orthogonal transformation, it becomes a trade off between the better fit to the inter-correlation matrix obtained by using an orthogonal transformation and the better fit to the rotated factor solution possible under certain circumstances with the orthonormal transformation. The differences between the two methods in regard to fitting the rotated factor tend to diminish as the number of factors involved increases. On the other hand, the advantage possessed by the orthogonally transformed solution in reproducing the inter-correlation matrix increases. Thus it is clear that the value of the orthonormal transformation as compared to the orthogonal transformation is least likely when the number of factors in the initial rotated solution is large. However, the number and nature of the non-overlapping variables in the two studies is also important.

Since the extension of the transformed solution to the non-common variables is an important aspect of these studies, the reproduction of the intercorrelations between the extended factors and these additional variables is an important consideration.

It is interesting that the advantage of the orthogonally transformed solution for reproducing the intercorrelation matrix did not always hold among the non-overlapping variables. This underlines the fact that the advantage of the orthogonal transformation matrix in reproducing the correlation matrix in the cross sample is partly due to its more efficient capitalization on sampling error. This sampling error effect is further underlined by the fact that for the orthonormal transformed solution, for one sample, the elements of the residual matrix involving the non-common variables were smaller than the elements of the corresponding matrix involving the common variables. This was, of course, just the opposite of the results obtained from the orthogonally transformed factor solution. However, while the initial unrotated solution extended to the non-common variables, necessarily possesses the maximization properties of the initial principal component solution for only the common variables, the advantage, while reduced, was still present for non-common variables in the larger samples.

The two following questions were raised at the conclusion of the two USAPRO presentations:

(1) Have factors (i.e., factor pure tests) proved to be good predictors of Army performance criteria?

(2) What is the advantage, for prediction, of using orthogonal predictors over the original correlated predictors if optimal weights are applied?

The two questions are closely related in that they are both concerned with the immediate application of factor analysis results to the practical problem of predicting personnel performance. USAPRO has considerable research evidence indicating that tests developed to measure factors do not predict performance as well as factorially complex tests developed to predict a specific Army performance measure. We have very little evidence bearing on our own factor measures, since, on theoretical grounds, we have not expected factor measures to have immediate use as predictors. Factors are useful constructs because of their simplified (i.e., more easily understood) relationships with psychological or physiological measures and the other factors. Thus, factor scores are useful in experimentally testing hypotheses relating carefully defined psychological content of a measure to human performance, and the factor concept has general usefulness for the better understanding of the psychological content of a battery of tests. It is not expected that factors will have immediate usefulness as operational predictors.

APPENDIX I

Formulae and Notation

I. Certain letters will be used consistently to denote specific kinds of matrices. Different matrices of the same type will be discriminated by their subscripts.

R a gramian matrix whose elements are product moment correlation coefficients.

P a principal component factor solution.

A an orthogonal eigen vector matrix derived from a gramian matrix.

D eigen value matrix.

F a factor solution other than a principal component factor solution.

T a transformation matrix whose elements are cosines of the angles between reference vectors (factors).

II. The following formulae relate several of the above matrices:

$$A'RA = D, \quad AD^{1/2} = P, \quad PP' = R, \quad P'P = D$$

$$FF' = R, \quad F'F \neq D$$

$$F_Y T_1 = F_X ; \quad T_1 = (F'_y F_y)^{-1} F'_y F_x$$

III. Table 1

A Sectioned Matrix Whose Elements are Projections*
Involving the Row and Column Variables

(As computed in the second sample)

		Rotated Factors from 1st Sample	PC Factors in Space Defined by Rotated Factors from 1st Sample	Experimental Variables
		$F_{x1} \cdots\cdots F_{xk}$	$F_{o1} \cdots\cdots F_{ok}$	$Y_1 \cdots\cdots Y_n$
Rotated Factors From First Sample ($k \leq n$)	$F_{x1} \vdots F_{xk}$	R_L $R_L = T_2' T_2$ $P_y T_2 = F_x$ $T_2 = D_y^{-\frac{1}{2}} A'$ **	P_{XL} $P_{xL} = A_L D_L^{\frac{1}{2}}$ $P_{xL} = R_L A_L D_L^{-\frac{1}{2}}$ $A_L' R_L A_L = D_L$	F_x'
Principal Component Factors (PC Factorization of R_y)	$P_{y1} \vdots P_{yn}$	T_2 $T_2 = D_y^{-\frac{1}{2}} A_y' F_x$ $T_2 = P_y^{-1} F_x$	T_3 $T_3 = T_2 A_L D_L^{-\frac{1}{2}}$	P_Y $P_y' = D_y^{\frac{1}{2}} A_y'$ $A_y' R_y A_y = D_y$
Experimental Variables	$Y_1 \vdots Y_n'$	F_X $F_x = F_y T_2$ $F_x = P_x A_x'$ $A_x(F_x' F_x) A_x = D_x$	F_{YL} $F_{yL} = F_y T_3$ $F_{yL} = F_x A_L D_L^{-\frac{1}{2}}$	R_Y

* These projections are cosines when both row and column vectors are of unit length.

** If $k = n$, $R_L = F_x' R_y^{-1} F_x$. However, in almost all practical situations, k, the number of rotated factors, will be considerably smaller than n, the number of variables.

IV. Green's Procedure

Problem: In order to fit a factor solution of k factors (i. e., P_{yk}) in the second sample to a rotated factor solution, F_x, in the first sample (i. e., $P_{yk}T_o = F_x$), compute T_o such that tr $(P_{yk}T_o - F_x)'(P_{yk}T_o - F_x)$ is minimized, while meeting the side constraint that $T_o' = T_o^{-1}$.

Solution:

(a)* $T_o = (P_{yk}' F_x F_x' P_{yk})^{-1/2} P_{yk}' F_x$, P_{yk} contains the k columns of P_y, the complete principal component factorization of R_y, corresponding to the k larger roots of D_y. $P_y = A_y D_y^{1/2}$ where $A_y' R_y A_y = D_y$ and R_y is the product moment correlation matrix for the experimental variables in the second sample.

(b) T_o can also be computed from the least square transformation,

$$T_1 = D_{yk}^{-1} P_{yk}' F_x$$

$$T_o = (D_y T_1 T_1' D_y)^{-1/2} D_y T_1$$

(c) A slightly different orthogonal transformation matrix, T_p, can be derived by directly orthogonalizing the T_1 as follows:

$$T_1'X = T_p' \ , \ T_p'X = T_1^{-1}$$

$$T_1'X = T_1^{-1} X^{-1} \ , \ X = (T_1 T_1')^{-1/2}$$

---

*Green, B. F. The orthogonal approximation of an oblique structure in factor analysis. Psychometrika, 1952, 17, 429-440.

$$T_p = (T_1 T_1')^{-1/2} T_1$$

Properties of T_o:

The orthogonally transformed factor solution, $P_{yk} T_o$, retains the maximum reproduction of R_y (i.e., $P_{yk} T_o T_o' P_{yk}' = P_{yk} P_{yk}'$). When communalities are substituted for ones in the diagonals of R_y and if the rank of R_y becomes K, this is undoubtedly the best procedure for fitting F_x.

V. Alternative Procedure

A non-square (orthonormal) transformation matrix permitting the full utilization of P_y can be developed as follows:

a. Whereas T_1 in the previous model provided a least square fit, T_2 in the model below provides an exact fit (since P_y has an inverse while P_{yk} does not when $k < n$.)

$$P_y T_2 = F_x ; \quad T_2 = P_y^{-1} \cdot F_x = D_y^{-1/2} \; A_y' F_x$$

b. The factor matrix F_x, computed and rotated in Sample 1, cannot usually be obtained by an orthogonal or orthonormal transformation of P_y, computed in Sample 2. The transformation of $P_y T_2$ into a solution within an orthogonal frame F_{yL}, can be accomplished as follows:

$$F_{yL} = P_y T_2 \; (A_L D_L^{-1/2})$$

c. While the matrix F_{yL} contains K column vectors (factors) spanning the same space as the oblique factors in F_x, the orthogonali-

zation was not accomplished in such a way as to maximize the fit of F_{yL} to F_x. The additional transformation required to effect this fit can be accomplished by using Green's procedure to obtain orthogonal matrix T_{o2} as follows:

$$F_{yL} T_{o2} = F_x$$

$$T_{o2} = (F'_{yL} F_x F'_x F_{yL})^{-1/2} F'_{yL} F_x$$

$$T_{o2} = \left[D_L^{-1/2} A'_L (F'_x F_x)^2 A_L D_L^{-1/2} \right]^{-1/2} D_L^{-1/2} A'_L (F'_x F_x)$$

d. Thus the ortnonormal transformation matrix T_m which minimizes the trace $(P_y T_m - F_x)'(P_y T_m - F_x)$, where T_m is an $n \times k$ orthonormal matrix is,

$$T_m = T_2 A_L D_L^{-1/2} \left[D_L^{-1/2} A'_L (F'_x F_x)^2 A_L D_L^{-1/2} \right]^{-1/2} D_L^{-1/2} A'_L (F'_x F_x)$$

VI. An orthonormal transformation matrix T_n providing a least square fit of T_n to T_2, as compared to the fit of F_{yL} to F_x in Part V, can be provided as follows:

a. The conversion of T_2 to an orthonormal matrix, T_3, spanning exactly the same space can be accomplished as follows:

$$T'_2 T_2 = R_L$$

$$T_3 = T_2 A_L D_L^{-1/2} \quad ; \quad T'_3 T_3 = I$$

b. While T_3 is an orthonormal matrix, it is not the orthonormal matrix with the best least square fit to T_2. There is still the need to minimize the trace of $(T_3T_{o2} - T_2)'(T_3T_{o2} - T_2)$, where T_{o2} is an orthogonal matrix. This can be accomplished by making use of Green's procedure described under IV(a).

$$T_{o2} = (T_3'T_2T_2'T_3)^{-1/2}T_3'T_2$$

$$T'_3T_2 = P^{-1}_{xL}T_2'T_2 = P^{-1}_{xL}R_L = P'_{xL} \; ; \quad \text{since } P^{-1}_{xL} = D_L^{-1/2}A_L'$$

$$T_{o_2} = (P'_{xL}\,P_{xL})^{-1/2}P'_{xL} = D_L^{-1/2}\;P'_{xL}$$

$$T_{o2} = A_L'$$

c. Thus the n x k orthonormal transformation matrix T_n which minimizes (in the least square sense) trace $(T_n - T_2)'$ $(T_n - T_2)$ is equal to,

$$T_3T_{o2} = T_2A_LD_L^{-1/2}A_L' = T_2\,(T'_2T_2)^{-1/2}$$

$$T_n = T_2\,(T'_2T_2)^{-1/2}$$

Note the similarity in the form of the computing formulae used to describe T_n and T_p. $T_p = (T_1T_1')^{-1/2}\,T_1$.

APPENDIX II

The Comparison of Transformation Matrices

I. Method

The variables included in the rotated factor solution in the first sample are designated by x and the rotated factor solution by F_x. The same (i.e., overlapping) variables in the second sample are designated by y and the non-overlapping variables by z. The factorization of R_y and R_z are accomplished as $P_y = R_y A_y D_y^{-1/2}$ and $F_z = R_{yz} A_y D_y^{-1/2}$. The transformed factor matrices in Sample 2 are $F_{yr} = P_y T$ and $F_{zr} = F_z T$. Each transformation matrix, T, computed by the methods described in Appendix I, is evaluated by determining the fit of F_{yr} to F_x and the reproduction of R_y by $F_{yr}F'_{yr}$, R_{yz} by $F_{zr}F'_{yr}$ and R_z by $F_{zr}F'_{zr}$. This is determined by comparing (for the different T matrices) the traces of the following product matrices:

$$(F_{yr} - F_x)'(F_{yr} - F_x) \;,\; (F_{zr}F'_{yr} - R_{yz})'(F_{zr}F'_{yr} - R_{yz})$$

$(F_{zr}F'_{zr} - R_z)'(F_{zr}F'_{zr} - R_z)$, and, after setting diagonal elements of $F_{yr}F'_{yr}$ and R_y equal to zero, $(F_{yr}F'_{yr} - R_y)'(F_{yr}F'_{yr} - R_y)$.

II. Results

The sums of squares of the residual matrices, computed as the traces of the matrices indicated in Part I above, are provided in Table 1 for a study involving physical proficiency measures. The x sample consisted of 254 West Point Cadets of the class of 1949. The y sample contained 294 West Point Cadets of the class of 1964. Table 2 relates to a study involving the following x and y variables (in samples 1 and 2 respectively): 15 "Personality" tests, 9 information tests, and 8 mental aptitude tests. Five rating variables based on performance as Infantryman, make up the z variables. Sample one (x variables) had 550 examinees and sample two(y variables) had 375 examinees.

The rank ordering of the magnitudes for the various entries in Tables 1 and 2 can be readily predicted from the algebraic formulations of the T's. The relatively efficiency for fitting F_x, going from high to low, is T_m, T_n, T_o, T_p. The relative efficiency for reproducing the R matrices is the same for all T's which are either orthogonal or capable of being linked by an orthogonal transformation. Thus all the orthogonal T matrices have more efficiency for reproducing R_y, when applied to PC solutions of R_y, than do the orthonormal transformations.

Table 1

Comparison of Transformation Matrices Computed on a Sample of 294 Examinees

(Physical Proficiency Variables)

Residual Matrices	Number of Elements Contributing to the Sums of Squares Reported as Entries in this Table	Total Sums of Squares of Elements in Residual Matrices						
		ORTHOGONAL Transformations				ORTHONORMAL Transformations		
		Communalities in Diagonals of R_y		Ones in Diagonals of R_y		Ones in Diagonals of R_y		Communalities in Diagonals of R_y
		T_o	T_p	T_o	T_p	T_m	T_n	T_m
$(F_{yr} - F_x)$	19 x 9 = 171	2.1381	3.1934	5.2665	5.7615	1.0943	1.5233	1.6758
$(F_{yr}F'_{yr} - R_y)$, other than diagonal elements	19 x 18 = 342	.2268	.2268	.8496	.8496	5.1851	5.1851	1.2382
$(F_{zr}F'_{yr} - R_{zy})$	19 x 11 = 209	.2953	.2953	.4129	.4129	3.0342	3.0342	.8050
$(F_{zr}F'_{zr} - R_z)$	$(11)^2$ = 121	2.1925	2.1925	3.6127	3.6127	6.2471	6.2471	2.3205

Table 2

Comparison of Transformation Matrices Computed on a Sample of 375 Examinees

(Personality, Mental Aptitude, Information, and Rating Variables)

Residual Matrices	Number of Elements Contributing to the Sums of Squares Reported as Entries in this Table	Total Sums of Squares of Elements in Residual Matrices						
		ORTHOGONAL Transformations				ORTHONORMAL Transformations		
		Communalities in Diagonals of R_y		Ones in Diagonals of R_y		Ones in Diagonals of R_y		Communalities in Diagonals of R_y
		T_o	T_p	T_o	T_p	T_n	T_m	T_m
$(F_{yr} - F_x)$	32 x 8 = 256	1.3258	1.3667	1.8747	1.9054	.2133	.1790	1.4270
$(F_{yr}F'_{yr} - R_y)$, other than diagonal elements	32 x 31 = 992	.6612	.6612	2.0005	2.0005	3.5448	3.5448	4.4173
$(F_{zr}F'_{yr} - R_{zy})$	32 x 5 = 160	.2162	.2162	.1965	.1965	.3405	.3405	.4163
$(F_{zr}F'_{zr} - R_z)$	$(5)^2$ = 25	8.898	8.898	9.1654	9.1654	9.8459	9.8459	8.0030